Ibrk at the NTCIR-14 QA Lab-PoliInfo Classification Task

Minoru Sasaki and Tetsuya Nogami
Ibaraki University

Introduction

- Stance Classification
  - automatically identify speaker's position on a specific target of topic from text.
  - The speaker's position is one of Three labels.
    - Support (favour/favor, agree, pro)
    - Against (oppose, disagree, con)
    - Neutral (none, unrelated, neither)
  - For example,
    - we want to know whether the former president Barack Obama is in favor of stricter gun laws from his speeches.

Introduction

- Previous researches have demonstrated many approaches to solve stance classification tasks.
  - (Rajadesingan 2014)
    - Use semi-supervised learning in online forum.
  - (Bamman 2015)
    - Use unsupervised method
  - (Ebrahimi 2016)
    - Use a supervised probabilistic classification in tweets.

Stance Classification Using Machine Learning

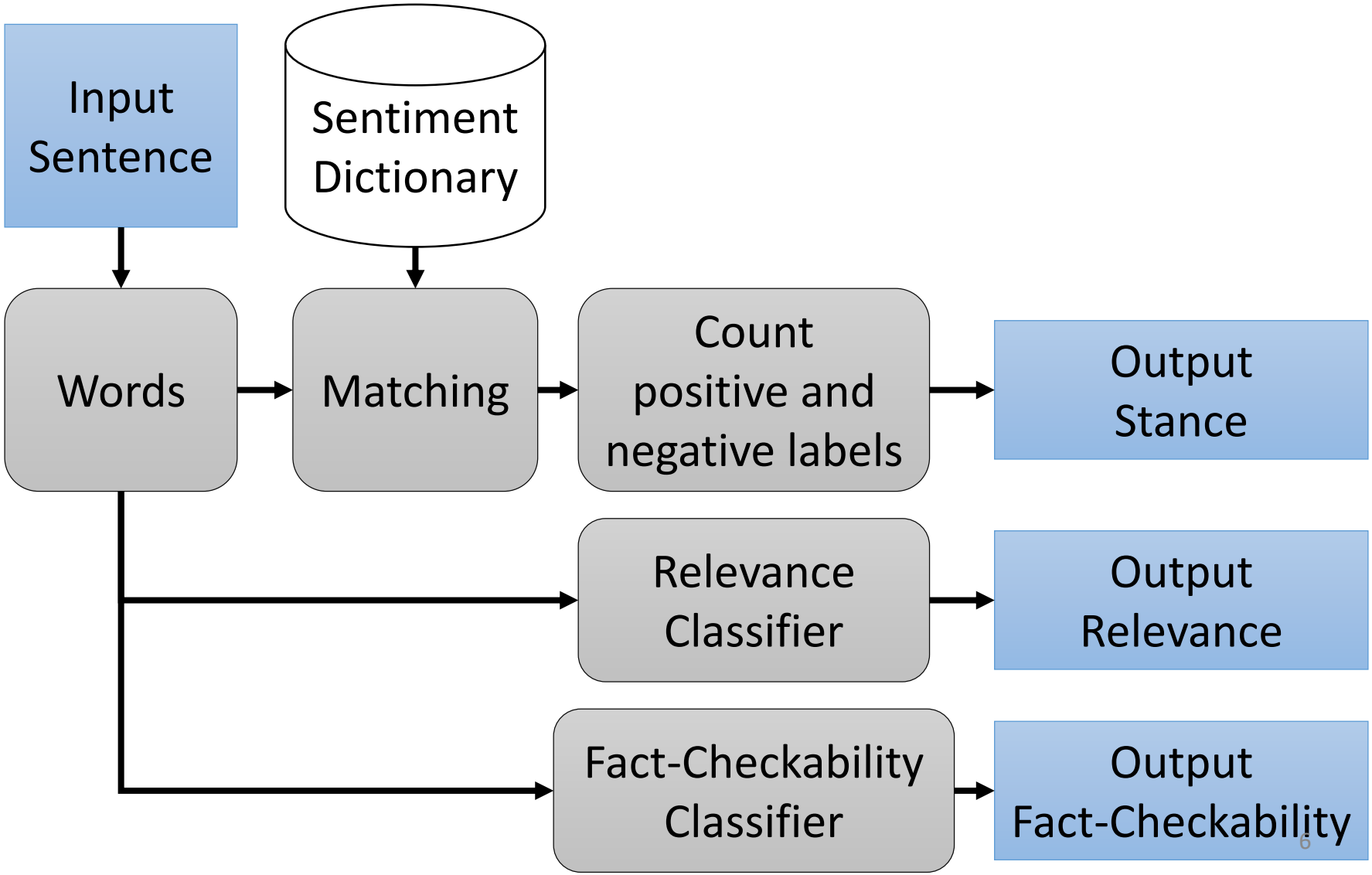
- In supervised approach,
  - this task is difficult due to imbalanced class sizes.
  - Stance classification task usually requires a large amount of training data to obtain many sentiment expressions.

- We propose to use sentiment dictionary for stance classification.
  - a sentiment dictionary is introduced to label each word with polarity information in the dictionary.

Purpose of This Study

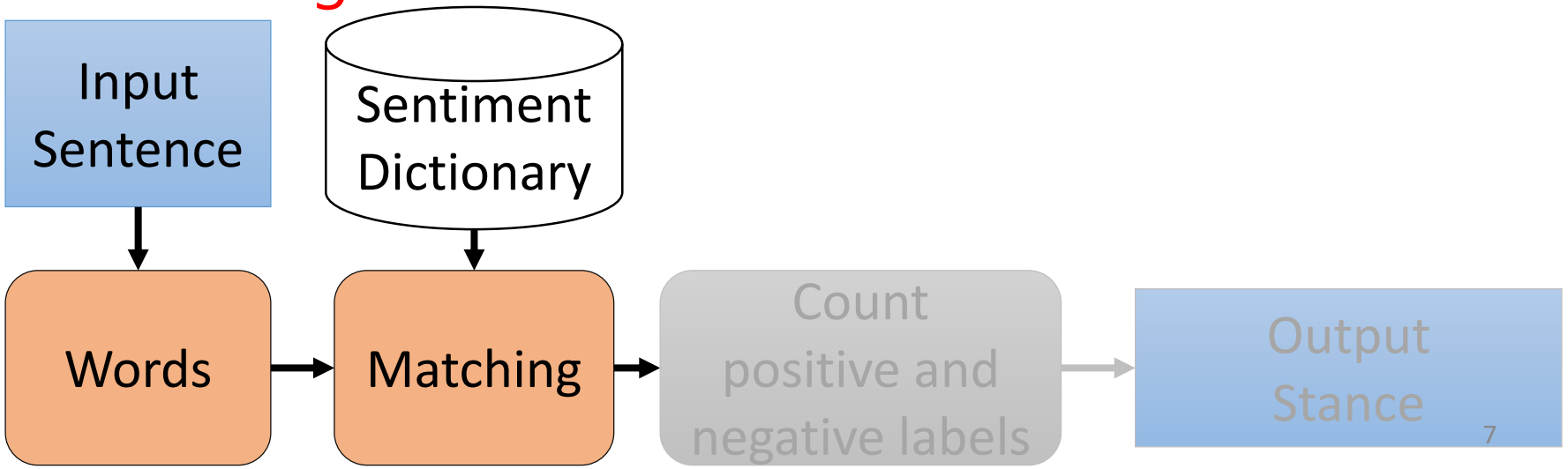
- We propose a stance classification system using sentiment dictionary.
- To evaluate the effectiveness of our system,
  - we conduct some experiments to compare with the result of the baseline method using Support Vector Machine (SVM).

System Description



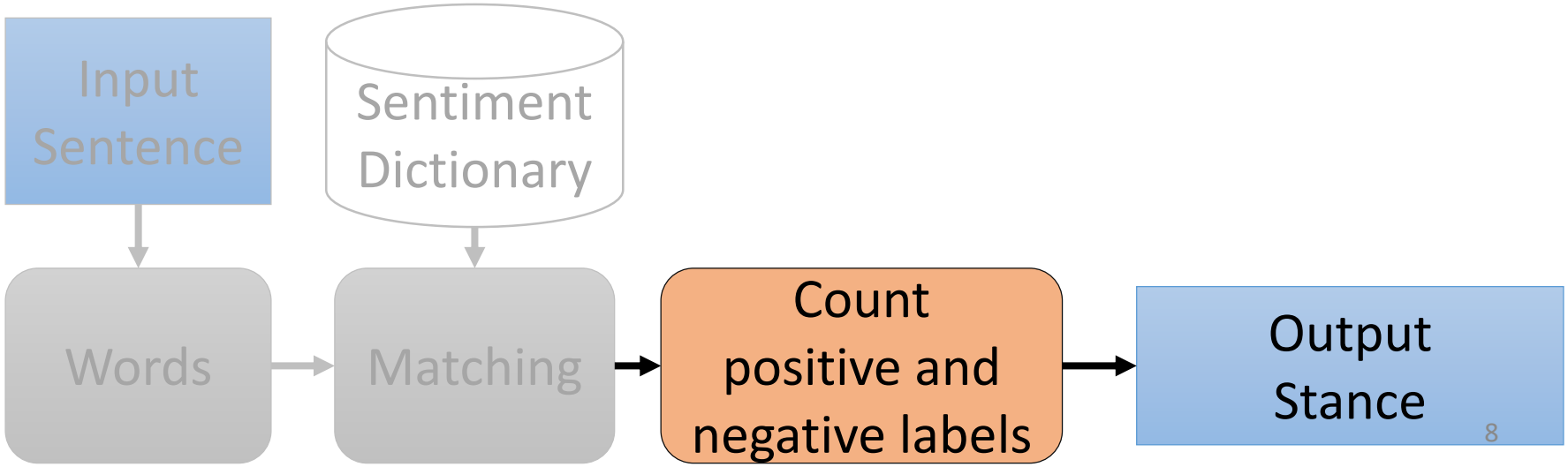
Stance Classifier (1/2)

- If each extracted word exists in the sentiment dictionary,
  - the polarity of the word is extracted to identify sentiment polarity label (positive or negative).
- The system counts up the number of positive and negative labels in the sentence.



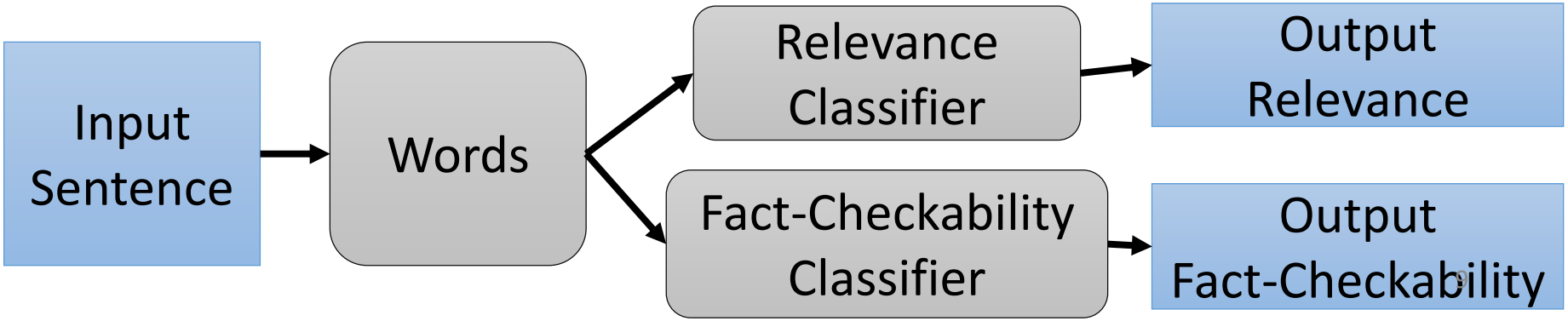
Stance Classifier (2/2)

- If the number of positive labels is greater than the number of negative labels,
  - the system assigns “support” label to the sentence, otherwise the system assigns “against” label.



Relevance Classifier and Fact-checkability Classifier

- We extract nouns, verbs and adjectives from the input sentence in the training data.
- Each set is represented as a feature vector by calculating frequencies of the features.
- We construct two classifiers by Support Vector Machine (SVM) from labeled feature vectors.
- The both classifiers are used to predict labels.



Experiments

- NTCIR14 QA Lab-PoliInfo Classification Task Dataset
  - 14 Topics
  - about 30,000 sentences in training data
  - 3,412 sentences in test data
- Sentiment Dictionary
  - Japanese Sentiment Polarity Dictionary
    - created by Tohoku University
  - We use this dictionary to obtain a sentiment polarity of word.

Experimental Results (1/6)

- Precision for the topic "Integrated Resort"

Methods	Support	Against	Neutral
Our System	**7.19%**	**15.63%**	**92.10%**
Baseline System	0%	0%	90.73%

- Precision, recall and F-measure for this topic

Methods	Precision	Recall	F-measure
Our System	77.80%	77.80%	77.80%
Baseline System	90.70%	90.70%	90.73%

Experimental Results (2/6)

- Precision for the topic “Integrated Resort”

Methods	Support	Against	Neutral
Our System	7.19%	15.63%	92.10%
Baseline System	0%	0%	90.73%

- The proposed system obtained higher precision than the baseline system using SVM.
  - These results show that the sentiment dictionary is effective for stance classification.
  - When we use the baseline system, all samples are classified into “neutral”.

Experimental Results (3/6)

- Precision, recall and F-measure of test data for this topic
  - All scores are decreased about 13% in comparison to the baseline system.
  - Because there are a lot of neutral samples in the training and test data.

Methods	Precision	Recall	F-measure
Our System	77.80%	77.80%	77.80%
Baseline System	90.70%	90.70%	90.73%

Experimental Results (4/6)

- Results for the “relevance” of the topic

label	Relevance		Not Relevance	
Method	Precision	Recall	Precision	Recall
Our System	86.50%	100%	NaN	0%

- All data were classified as relevant to the topic.
  - It is difficult to detect sentences that are not related to the topic by using SVM.

Experimental Results (5/6)

- Results for the “fact-checkability” classification

label	fact-checkable		not fact-checkable	
Method	Precision	Recall	Precision	Recall
Our System	NaN	0%	64.6%	100%

- All data were classified as “not fact-checkable”.
  - It is difficult to detect sentences that we can conduct a fact-check by using SVM.

Experimental Results (6/6)

- Results for the class label using our system

label	Precision	Recall	F-measure
fact-check-support	6.3%	17.8%	9.3%
fact-check-against	4.5%	20.2%	7.4%
class-other	93.4%	77.0%	84.4%

- The small number of test data can be classified correctly.
  - In the future, we will improve our system to classify "class-other" samples effectively.

Conclusions

- We proposed a new method for stance classification using sentiment dictionary.
- The effectiveness of the proposed method was evaluated on the NTCIR-14 QA Lab-PoliInfo classification task formal run dataset.
- The experimental results show that the proposed methods obtains higher precision than the baseline method using SVM.
  - However, the precision of our system is decreased about 13% in comparison to the baseline system for the “neutral” samples.